Optimum Inductive Methods. A study in Inductive Probability, Bayesian Statistics, and Verisimilitude.

Festa, Roberto

SUMMARY

According to the Bayesian view, scientific hypotheses must be appraised in terms of their posterior probabilities relative to the available experimental data. Such posterior probabilities are derived from the prior probabilities of the hypotheses by applying Bayes'theorem. One of the most important problems arising within the Bayesian approach to scientific methodology is *the choice of prior probabilities*. Here this problem is considered in detail w.r.t. two applications of the Bayesian approach: (1) the theory of inductive probabilities (TIP) developed by Rudolf Carnap and other epistemologists and (2) the analysis of the multinomial inferences provided by Bayesian statistics (BS).

The subjective view and the aprioristic view represent the two 'traditional' views regarding the choice of prior probabilities. According to the subjective view such a choice is restricted only by the axioms of the probability calculus. On the contrary, according to the aprioristic view, the choice is controlled by further a priori principles of rationality. In this book a critical analysis of the traditional views is provided and a different view - called the *contextual view* - is proposed (Chapter 7). According to the contextual view the selection of the optimum prior probabilities is restricted by the cognitive context of a given empirical inquiry. More precisely, two kinds of contextual constraints are considered viz. the background knowledge shared by the scientists engaged in the inquiry and the cognitive goals pursued by these scientists. In particular, it is assumed that the main cognitive goal of science is the achievement of a high degree of verisimilitude or approximation to the truth. This claim - called the verisimilitude thesis - is a basic assumption of the current epistemological verisimilitude theory (VT) developed by Sir Karl Popper and others.

The Bayesian approach and VT belong to two different fallibilistic methodological traditions which originated as a result of the crisis of infallibilism (Chapter 1.2). According to infallibilism the cognitive goal of science is certainty regarding the truth of the adopted theories. From antiquity to the first half of the eighteenth century, infallibilism was the prevailing methodological view in the philosophy of science. However, from the second half of the seventeenth century, it became increasingly clear that infallibilism was indefensible and that more modest 'methodological ideals' were needed. Hence two different fallibilistic views of scientific knowledge were proposed: the probabilistic view and the verisimilitude view. The probabilistic view suggests that in most cases knowledge is merely probable (contrary to the infallibilistic tenet that certainty is the only epistemic state compatible with science). On the other hand, the verisimilitude view suggests that science, in its development, aims to move closer and closer to the truth (contrary to the infallibilistic tenet that science has to reach the exact truth).

Although the probabilistic view and the verisimilitude view evolved into two distinct methodological traditions often in competition with each other, the present inquiry is inspired by the conviction that a 'mature' fallibilistic methodology can be developed by integrating the two traditions. The contextual view of prior probabilities proposed in this book can be seen as a step in this direction since it provides an example of 'co-ordinated development' of TIP, BS and VT where TIP and BS belong to the tradition of probabilistic fallibilism and VT belongs to the tradition of verisimilitude fallibilism.

In inductive inferences the informative content of the conclusions (hypotheses) is not entirely contained in that of the premises (evidence): hence one can never be completely sure that hypotheses are true. According to the Bayesian view of inductive inference the degree of uncertainty of an individual regarding the hypotheses can be formally represented by appropriate epistemic probabilities. In particular, BS aims to provide a systematic analysis of statistical inferences. Furthermore, TIP offers a Bayesian analysis of certain types of inductive inferences, such as prediction of future events, which are also typical subjects of study in philosophical research on induction.

Both BS and TIP can be used to analyze certain inductive inferences - called multinomial inferences - relative to a multivariate Bernoulli process *Ex* whose trials are classified into a set $\mathbf{Q}$ formed by k mutually exclusive and jointly exhaustive categories, or properties, $Q_1,\ldots,Q_k$. In BS the analysis of multinomial inferences is made by using certain prior probabilities, or *prior distributions*, defined on the possible values of the parameter vector $\mathbf{q}=(\mathbf{q}_1,\ldots,\mathbf{q}_k)$ where $\mathbf{q}_i$ denotes the objective probability of obtaining Q_i (with $i = 1,\ldots,k$) in a trial of *Ex*. On the other hand, in TIP the analysis of multinomial inferences is made by using appropriate *inductive methods* i.e appropriate predictive probabilities defined on the possible results of certain sequences of trials of *Ex*.

In spite of the different strategies used by TIP and BS there is a strict relationship between these two approaches to multinomial inferences. Indeed it can be proved, by using de Finetti's representation theorem, that any inductive method belonging to the class of the so-called exchangeable inductive methods is 'equivalent' to a unique prior distribution on $\mathbf{q}$ (Chapter 5.1). In particular the well known inductive methods proposed by Carnap and Stegmüller (1959) - herein referred to as GC-systems - are equivalent to the Dirichlet distributions commonly used in BS (Chapter 6.3).

The problem of choosing the optimum GC-system from a given set of GC-systems - or, equivalently, the optimum Dirichlet distribution from a given set of Dirichlet distributions - is the main technical problem considered in this book (Chapter 8). This problem - which is referred to as *the epistemic problem of optimality* (*EPO*) - can be stated as follows. A GC-system is fully characterized by a couple (γ,λ) where γ is the so-called prior vector and λ is a parameter included in the range $[0,\infty]$. Assuming that the prior vector $\gamma°$ has been selected as the 'optimum' prior vector for an inquiry into a given multivariate Bernoulli process *Ex*, then it may be asked which is the optimum GC-system $(\gamma°,\lambda°)$ within the set of all the GC-systems $(\gamma°,\lambda)$ with prior vector $\gamma°$ or, equivalently, which is the optimum λ-value $\lambda°$?

The proposed contextual solution of EPO (Chapter 8.3) is based on the following two assumptions:

(1) the background knowledge shared by the scientists engaged in the inquiry into *Ex* includes an 'informal estimate' of the Gini diversity $G(\mathbf{q}) = 1 - \Sigma \mathbf{q}_i^2$ of *Ex*, where Gini diversity is a measure of the degree of disorder of a population or process;

(2) the cognitive aim pursued by the scientists is to achieve a high degree of verisimilitude or, more precisely, to minimize the distance between their $(\gamma°,\lambda)$-based estimates of $\mathbf{q}$ and the truth (the true value of $\mathbf{q}$).

The prospective audience of this book includes readers interested in: (1) the theory of inductive probabilities (TIP); (2) the problem of selection of prior probabilities within Bayesian statistics (BS); and (3) the verisimilitude theory (VT). More generally, it includes anyone interested in statistics and the philosophy of science. Given that many of such readers may be relatively unfamiliar with any of the three above mentioned subjects, the first part of the book (Chapters 2-4) provides introductory material on TIP, BS and VT.

In the second part (Chapters 5-6) the basic features of GC-systems and Dirichlet distributions, and the relationships between the two, are illustrated.

Finally, in the third part (Chapters 7-11) the contextual view of prior probabilities is illustrated. In particular, in Chapter 7.6 a context-dependent justification of the decision to use a prior Dirichlet distribution within a given inquiry is provided. The above mentioned contextual solution to EPO is presented in Chapter 8. Another technical result presented in this chapter (Section 4) is the solution of the so-called *logical problem of optimality* for GC-systems. In Chapter 8.7 it is proved that this solution generalizes a well-known result obtained by Carnap (1952). In Chapter 9 some features of the contextual view of prior probabilities are elucidated by comparing this view with other approaches to EPO. In Chapter 10 Gini diversity, which plays a crucial role in the contextual approach to EPO, is considered in some detail. Lastly, in Chapter 11, the relationships between the contextual view and some recent methodological programmes of research are considered and some possible developments of the contextual view are suggested.